Australian Government
Department of Defence
Defence Science and
Technology Organisation

ACSPRI 2014 4th International Social Science Methodology Conference Report

Elena Mazourenko

Joint and Operations Analysis Division
Defence Science and Technology Organisation

DSTO-GD-0878

ABSTRACT

This report summarises themes of interest to the Technology Forecasting and Futures (TFF) Group of JOAD presented at the ACSPRI 4th International Social Science Methodology conference. These themes include: causal inferences in quialitative research; the role of theory and interpretation in analysing Big Data; the role of paradata in increasing data quality; the Total Survey Error framework; multi-modal on-line surveying, quality frameworks for assessing qualitative research; and the dynamics of manual vs automated data analysis. Research findings related to these themes are contextualised for the TFF S&T forecasting methodology leading to series of recommendations for improving the S&T forecasting practice.

Published by

Joint and Operations Analysis Division
DSTO Defence Science and Technology Organisation
PO Box 1500
Edinburgh South Australia 5111 Australia

Telephone: 1300 333 362
Fax: (08) 7389 6567

AR-016-331
April 2015

APPROVED FOR PUBLIC RELEASE

ACSPRI 2014 4th International Social Science Methodology Conference Report

Executive Summary

The Australian Consortium for Social and Political Research Incorporated (ACSPRI) hosted the 4th International Social Science Methodology at the University of Sydney on 7-10 December 2014. The conference attracted a vibrant international community of researchers, academics, and practitioners in public and private sectors from Australia, New Zealand, USA, Malaysia, Canada, Sweden, UK, and China. The goal of the conference was to provide a forum for discussions and debate on current and emerging issues in social science methodology, and to present best practice standards and innovation within a range of disciplines.

A member of the Technology Forecasting and Futures (TFF) Group of the Joint Operations Analysis Division (JOAD) attended the conference. This report covers the themes of interest to the TFF's S&T forecasting practice presented at the conference followed by series of recommendations to the TFF Group and broader DSTO community involved in technological forecasting.

A conference workshop titled 'Can qualitative approaches support causal inferences? Emerging design options and analytic techniques' provided an overview of the current perspectives on causal claims in qualitative research. Three approaches to generating plausible causal relationships were discussed: Qualitative Comparative Analysis (QCA), Process Tracing, and Program Theory. QCA is applicable to the TFF's forecasting practice for cross-case comparisons to identify typological patterns for medium to long term forecasting. Process Tracing is also applicable to medium and long term forecasting. It is based on linking and analysing putative causes to observed effects and includes analysis of 'rival' explanations of phenomena. Program Theory is useful as a method designed to highlight causal inferences in studies and projects. This is achieved by theory building and testing within the context of each application.

The role of theory and interpretation in extracting and analysing 'useful' data from the torrent of Big Data is continuously growing. Multi-disciplinary approaches with a mixture of research traditions and methods reflect the current stance of social researchers when 'dealing' with Big Data. Multi-disciplinary analytical toolkits are investigated and applied by the TFF group as part of the S&T forecasting practice.

Social researchers apply 'responsive designs' to data collection processes for effective use of paradata (information about the process of collecting survey data) to improve survey outcomes. Paradata are monitored by a production monitoring dashboard, which uses information about data collection processes to help guide alterations in field protocols leading to greater efficiency and improvements in data quality. 'Quality

indicators' applied to analysing paradata in recent studies can be applied to both automated and manual data collection as part of TFF studies. Published studies can also provide useful insights to various approaches that can be taken to reduce error in survey design and conduct.

Total Survey Error (TSE) is a term used to refer to all sources of bias (systematic error) and variance (random error) that may affect the quality of survey data. The TSE framework, presented at the conference, can be used for effective planning of research, guiding decision making about data collection, and contextualising the interpretation and dissemination of findings. Furthermore, this approach can optimise survey design within given research constraints. The usefulness of the TSE framework for technological forecasting should be investigated. Of particular interest to the TFF studies are the dimensions of data accuracy, credibility, and relevance.

There is a rapid increase in the use of mobile devices such as tablets and smartphones in conducting on-line surveys. However, survey methodologies often fall behind in providing survey development tools to address multi-modal data collection. Recent research identified a link between these shortcomings and reduced response rates and accuracy of results. Published studies showed that appropriate design modifications can lead to minimising the difference is response rates and bias in survey data. These findings are useful for designing TFF studies, which also include multi-modal data collection.

The notion of quality in qualitative research has been long debated by researchers. There is a shared commitment to rigour, robustness and relevance, but without a solid agreement on what this means or to what extent these measures can be formalised. The outcomes of this debate to date were discussed at the conference. Central to this discussion were two approaches: a framework for assessing research evidence and a quality framework based on transparency and systematicity. The first approach utilises eighteen appraisal questions, each supported by quality indicators, while the second provides a choice of techniques to establish rigour at each stage within the qualitative research process. These frameworks, complemented by a detailed reflective research diary, can be utilised by the TFF group for assessing the quality of forecasting studies.

On the basis of findings described above, it is recommended that:

- Quality frameworks are tailored to the objectives of the TFF's Code of Best Practice and assessed for their usefulness to forecasting studies
- The role of theory and interpretation in extracting 'data of interest' from the Big Data is investigated within the context of forecasting studies. Multi-disciplinary approach to this investigation can be supported by DSTO's Communities of Practice
- The role of paradata in increasing data quality is assessed within forecasting studies including appropriate modifications to survey design, conduct, and analysis
- TSE framework is assessed for its usefulness to reducing survey error in forecasting studies;
- The findings and recommendations of published studies related to multi-modal on-line data collection are used for designing TFF surveys

- Manual and automated data analysis techniques as complementary research mechanisms are proposed to the DSTO Communities of Practice as an 'enduring theme' to facilitate on-going exchange between multi-disciplinary teams and learn from each other.

Author

Dr Elena Mazourenko
Joint and Operations Analysis Division

Elena joined DSTO in 2010 during the final stages of her PhD candidature with the Flinders University, which she completed successfully in 2011. She worked on a range of projects including the development of supplements for staff training programs at HQJOC. Elena worked in DSTO's support to operations area in 2011, during which time she was awarded a Secretary of Defence Fellowship to conduct a research into the application of Program Theory to support Defence planning and assessment processes. Upon completion of the Fellowship in 2013, she started working for the Joint Innovation Centre (renamed in 2014 to Technology Forecasting and Futured Group) of JOAD. Elena's research interests include the development and refinement of methods for technology assessment and S&T forecasting.

Contents

Figures

This page is intentionally blank

1. Introduction

The purpose of this report is to summarise the information presented at the Australian Consortium for Social and Political Research Incorporated (ACSPRI) 4th International Social Science Methodology Conference and identify its practical value to the Technology Forecasting and Futures (TFF) Group's S&T Forecasting practice. The themes of interest to the TFF's practice included: the emerging trends in increasing the capacity of qualitative research approaches to support causal inferences; a proposition that in social science context theory and models are required to identify correlations (and causality) in Big Data; innovations in automated data collection; measurement and other errors in data collection (with focus on Total Survey Error); recent developments in on-line surveying; and measures for increasing the quality in qualitative research.

The themes of interest are described in this report followed by a discussion on how the latest developments in these areas can influence, or improve, the S&T Forecasting practice. The report concludes with recommendations to the TFF Group and broader DSTO community involved in technological forecasting.

2. Background

The ACSPRI have hosted several Social Science Methodology conferences since 2006. The goal of these conferences is to provide a forum for discussions and debate on current and emerging issues in social science methodology to the national and international community of researchers and academics. The multi-disciplinary dimension of the ACSPRI conferences provides an opportunity to stay abreast of best practice standards and innovation within a range of disciplines, and explore opportunities for collaboration between social science researchers.

The 4th International ACSPRI Social Science Methodology Conference was held at the University of Sydney on 7-10 December 2014. It attracted a vibrant international community of researchers, academics, and practitioners in public and private sectors from Australia, New Zealand, USA, Malaysia, Canada, Sweden, UK, and China. The conference program included keynotes, plenary and paper presentation sessions, and workshops. The richness and diversity of the conference program provided the delegates with multiple choices for attending sessions of highest value to them and the organisations they represented.

3. Themes of interest to TFF

3.1 Causality in qualitative research

There is a growing interest among social researchers about the capacity of qualitative research to support causal analysis. The debates in the social research community reflect differing conceptions of approaches that underpin causality. In quantitative design and analysis procedures, the focus is normally on determining the extent to which the variance in one dimension causes variance in another. In contrast, qualitative design and analysis procedures are likely to be focused on understanding the role of one dimension in causing or influencing another dimension (Goodrick, 2014).

A conference workshop titled 'Can qualitative approaches support causal inferences? Emerging design options and analytic techniques' provided an overview of the current perspectives on causal claims in qualitative research. In particular, the Qualitative Comparative Analysis and Process Tracing were identified as techniques that have been successfully applied and proved useful for supporting causal analysis in qualitative studies. The workshop facilitator drew from a number of publications in this field to illustrate a promising design strategy for generating plausible causal claims. It was proposed at the workshop that these techniques offer robust approaches to qualitative data analysis.

Qualitative Comparative Analysis (QCA) is a technique applied within and across cases to develop and test causal pathways. The central goal of QCA is an exhaustive explanation of the phenomenon under investigation (Legewie, 2013). An underlying assumption of QCA is that social phenomena involve complex causality (Legewie, 2013). QCA provides a boost in analytic potential for cross-case comparisons and is particularly useful for medium-N data sets. It helps to make research more systematic and transparent and provides insights into causal and typological patterns that assist in developing mid-range theories (Legewie, 2013).

The Process Tracing technique is based on using evidence from within case based studies to make causal inferences. The general method of process tracing is to generate and analyse data on the causal mechanisms, or processes, events, actions, expectations, and other intervening variables, that link putative causes to observed effects in an attempt to determine the plausible explanations of phenomena (Bennett & George, 1997). Rival explanations are also identified, analysed, and eliminated.

In addition, Program Theory (Funnell & Rogers, 2011) was discussed at the workshop as a powerful method designed to highlight causal inferences in studies and projects. Theory building and theory testing are conducted within a context of each application. In general, context is seen within the research community as central to understanding causation in qualitative studies. Program Theory has a track record of successful applications within the field of program evaluation.

It was concluded at the workshop that while the capacity of qualitative research to defend causal inferences remains a debated issue within the research community, the continuous methodological work and outcomes of many studies suggest that this capacity is on a steady increase.

3.2 Big Data: challenges and advantages

Social scientists are increasingly using large-scale datasets from the Web to seek answers to long-standing questions about social, economic and political behaviour (Ackland, 2014). Big Data and in particular, social media data, present both methodological challenges and opportunities in empirical social science research. There are two prevailing views on how Big Data can transform social science. One view is that theory and interpretation will become less necessary as data will 'speak for themselves'. A counter view argued at the conference (Ackland, 2014) is that in social science context, theory and models are required to identify meaningful correlations and causality in Big Data. Furthermore, it is proposed that theory and interpretation are more necessary than before to find the appropriate layer of information in what otherwise is an 'unnavigable sea' of data (Gonzalez-Bailon, 2013).

As a 'network of networks' and the 'internet of things', Big Data represents the web of information that is available, but not necessarily easily obtainable or understood. Therefore, Big Data creates imperative to innovate methods for more effective 'management' of information of interest and value to scientists and researchers. Virtual worlds offer good-quality, time-stamped, micro-level data on social networks in large, heterogeneous populations (Burt, 2011). However, a construct-validity question has to be answered before research can translate between virtual worlds and real world (Burt, 2011).

In social sciences (SS), 'ground truth' is hard to establish. The rapidly growing role of Big Data in SS research is being investigated and published in journals such as 'Big Data and Society'. Currently, scientists apply 'traditional' SS statistical techniques such as regression using survey data. For example, statistical network analysis (e.g. ERGM) is derived from logistic regression (Ackland, 2014).

Big Data research can transform policy making by improving communication and governance in policy-making domains (Gonzalez-Bailon, 2013). However, there is an implicit assumption that the analyses of large datasets are not the dominion of a single discipline or approach, but a mixture of research traditions and methods within multi-disciplinary efforts (Gonzalez-Bailon, 2013).

The Virtual Observatory for the Study of Online Networks (VOSON) was established in 2005 at the Australian National University (http://voson.anu.edu.au). VOSON aims to advance the social science of the Internet by conducting research, developing research tools, and providing research training. The VOSON software for hyperlink network construction and analysis has been publicly available since 2006 and has been used by over 1500 researchers worldwide (Ackland, 2014).

3.3 Automated data collection

Current work in 'responsive designs' for data collection incorporates design and management strategies for effective use of paradata in order to improve survey outcomes (Wagner, West, Kirgis, Lepkowski, Axinn, & Ndiaye, 2012). 'Responsive designs' allow for the conceptualisation and implementation of design features that respond to survey conditions in real time (Heeringa & Groves, 2006). This is achieved by extensive use of paradata from a sample management system organised into a production monitoring dashboard, which uses information about data collection field work to help guide alterations in field protocols during survey data collection leading to greater efficiency and improvements in data quality (Kirgis, 2014).

In addition, paradata from audit trails can be used for data quality monitoring at the interviewer level. In recent years, the interviewer-level data quality dashboards for collecting high quality data have been implemented in several studies (Kirgis, 2014). Paradata are used to monitor interviewer's behaviour as part of survey quality control (Lin, Kelley, Mneimneh, & Pennell, 2014). The 'traditional' quality control procedures such as calling back respondents, audio-recording interviews, accompanying interviewer and observing the interview process can now be supplemented with the use of paradata. The key 'performance indicators' include the average time spent on survey questions, the frequency of using help screens, recording remarks, checking errors, backing up in the interview, and the frequency of 'don't know' and 'refuse' responses.

In recently conducted Saudi National Health and Stress Survey (SNMHS) (Lin, Kelley, Mneimneh, & Pennell, 2014) fieldwork was centrally monitored by the Survey Research Centre at the University of Michigan. Interviews were conducted face-to-face using a computer assisted personal interviewing (CAPI) mode, which collected paradata to monitor interviewer's behaviour. Quality indicators used in this study included time spent on survey questions, pauses taken by interviewers during the interview, interview length, verification results, response rate, eligibility rate, number of contact attempts, time between interviews, and a number of completed interviews per day. The SNMHS study provided an example of how making consolidated tools available at the beginning of data collection and coordinated by a remote organisation utilised paradata to allow real-time monitoring and control of data collection in dispersed and geographically distant field operations (Lin, Kelley, Mneimneh, & Pennell, 2014).

3.4 Errors in data collection

Total Survey Error (TSE) is a term used to refer to all sources of bias (systematic error) and variance (random error) that may affect the quality of survey data (Pennay, 2014). The TSE provides a framework (Figure 1) that supports the effective planning of research, guides decision making about data collection, and contextualises the interpretation and dissemination of findings (Whiteley, 2014). The TSE framework also allows researchers to systematically evaluate and improve the design and execution of ongoing survey programs and future investigations (Whiteley, 2014).

The TSE paradigm is part of a broader concept of Total Survey Quality (TSQ), which introduces other dimensions important to data users such as accuracy, credibility, comparability, usability, relevance, accessibility, completeness, coherence, and timeliness. Many statistical agencies, such as Australian Bureau of Statistics, have TSQ frameworks that guide their overall approach to survey research (Pennay, 2014).

The TSE paradigm relates to making survey design decisions and trade-offs to allocate resources in a way that reduces TSE for key estimates, thereby *optimising* survey design within given resource constraints (Pennay, 2014).

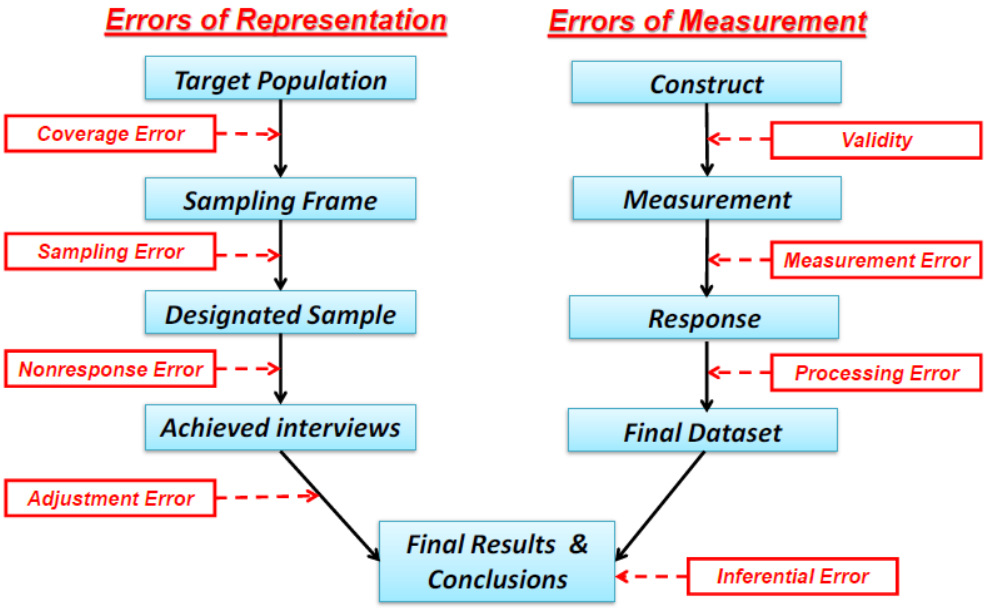

Figure 1 Total Survey Error Framework (Pennay, 2014)

The survey research literature contains many findings related to costs, errors and methodological effects of survey design. These findings are often generalisable to similar studies (Biemer, 2010). However, researchers can conduct their own experiments aimed at reducing TSE. The best way to undertake this type of experimentation is to derive some proxy indicators of data quality for their surveys and then measure the impact of alternative designs on data quality (Pennay, 2014).

Figure 2 summarises the identified strengths and weaknesses of the TSE approach. Applying this approach to survey design, implementation and evaluation can assist researchers in a number of ways. The TSE framework (Pennay, 2014):

- Provides theoretical and practical framework for all aspects of survey design and evaluation
- Enables researchers to challenge accepted paradigms regarding the primacy of response rates as an indicator of survey quality
- Helps guide survey design decisions
- Can be used as an organising framework for proposals and technical reports

- Is a tool for evaluating survey designs, helping to make informed decisions and driving continuous improvement
- Informs resource allocation because optimal research design equates with value for money.

Weaknesses	Strengths
Despite being around since 1944 (Denning) it has not become the dominant paradigm for survey researchers.	Provides a theoretical and practical framework for survey methodologists
Total MSE can rarely be completely measured which makes fully implementing a TSE approach challenging	Decomposition of errors and separation of issues
Survey researchers remain primarily focused on sampling errors	Increases the focus on non-sampling errors
The exclusion of key quality concepts found in overarching total quality frameworks	Increases the focus on achieving optimal design outcomes
	Makes explicit what is otherwise implicit
	Can be adapted for all forms of social, behavioural and market research (including qualitative research).

Figure 2 Strengths and weaknesses of TSE (Groves & Lyberg, 2010)

3.5 On-line surveying

Experts predict that mobile Internet usage will overtake desktop Internet usage worldwide by 2015 (Cobb, 2014). There is a rapid increase in the use of mobile devices such as smartphones and tablets for completing web surveys. Survey methodologies, however, are often behind in providing new survey development tools for conducting mixed-mode web surveys. There is an emerging empirical evidence (Cobb, 2014) of the implications of ignoring mode differences in the design of web survey on response rates and accuracy of results. Current research offers design options for dual-mode (desktop and mobile) optimisation of web surveys.

Mobile web surveys have distinct features such as small screens, smaller keyboards, and different navigation (Lattery, Park Bartolone, & Saunders, 2013). There are different 'techniques' in using desktop/laptop and mobile devices including use of mouse vs. touch screens, typing with thumbs on mobile devices, etc. In addition, respondents on mobile devices may be in various locations or situations when accessing surveys and therefore, may be distracted or have less time than a typical web-based survey completed on a desktop PC. All these factors impact various aspects of the response process and should be taken into consideration when designing a dual-mode survey (Lattery, Park Bartolone, & Saunders, 2013).

A study by (Lattery, Park Bartolone, & Saunders, 2013) tested five design enhancements to measure and compare completion rate, bias, and respondents' self-reported satisfaction in a dual-mode survey. Mobile users were auto detected and given a mobile rendered version of the survey with larger font designed for smartphone devices. The design enhancements included: number of scale points, grid format, grid length, radio button spacing, and background image.

The study results highlighted the importance of design modifications in dual-mode survey to minimise difference is response rates and bias. For example, imagery used on mobile devices significantly reduced response rate and increased time taken to complete the survey. While current body of research into the dual-mode survey designs is limited, researchers should utilise findings of available studies when designing their own, and make the results of their work available for the benefit of others.

3.6 Quality in qualitative research

There is no unified body of theory, methodology, or method that can collectively be described as qualitative research; therefore any attempt to establish a consensus on quality criteria for qualitative research is unlikely to succeed (Rolfe, 2006). Researchers have long debated the notion of quality – how it is defined and measured – and have demonstrated a shared commitment to rigour, robustness and relevance, but without a solid agreement on what this means or to what extent these measures can be formalised (Kellard, 2014).

The purpose of the conference stream concerned with the quality in qualitative research was to present and discuss the outcomes of this debate thus far. The materials for discussion were drawn from the progress made in the UK, USA and Canada over the last decade.

There are three distinct positions in the ongoing debate (Hope & Waterman, 2003): first, the adoption of positivist validity criteria for qualitative research; second, the establishment of distinct and separate criteria from those adopted in quantitative research; and third, a complete rejection of all predetermined criteria.

At the conference, the discussion was centered on the second position. The focus was on factors affecting quality in qualitative studies. It was proposed that there are four 'guiding principles' (Kellard, 2014), (Spencer, Ritchie, Lewis, & Dillon, 2003) of qualitative research:

- It adds new knowledge or evidence
- The research design is defensible
- The research conduct is rigorous
- The findings are credible and directly linked to data.

Can this study withstand external scrutiny? Are the findings valid? Is the new knowledge generated? These are the questions the researchers should be asking themselves when designing and conducting qualitative studies. 'Quality checks' are necessary at all stages of the research process: design, sampling, recruitment, fieldwork, analysis, and reporting. What indicators should be used and why? How do they differ in face-to-face, telephone

and online settings? There are also quality elements that run throughout the research process: research ethics, neutrality, reflexivity, appropriate researcher skills, etc.

There are several common features related to quality of qualitative studies described in the literature: the need for clarity in aims and objectives; appropriate use of qualitative methods; appropriate sample design; clarity about the analytical process; and clarity about how the evidence and conclusions are derived (Spencer, Ritchie, Lewis, & Dillon, 2003). Some authors (Rolfe, 2006) recommend that a detailed reflexive research diary is included with every research report in order for their quality to be properly assessed by a reader.

A framework for assessing research evidence offered by (Spencer, Ritchie, Lewis, & Dillon, 2003) contains eighteen appraisal questions, each supported by quality indicators:

1. How credible are the findings?
2. How has knowledge or understanding been extended by the research?
3. How well does the evaluation address its original aims and purpose?
4. How well is the scope for drawing wider inference explained?
5. How clear is the basis of evaluative appraisal?
6. How defensible is the research design?
7. How well defended are the sample design/target selection of cases/documents?
8. How well is the eventual sample composition and coverage described?
9. How well was the data collection carried out?
10. How well has the appraisal to, and formulation of, analysis been conveyed?
11. How well are the contexts of data sources retained and portrayed?
12. How well has diversity of perspective and content been explored?
13. How well has detail (i.e. richness) of the data been conveyed?
14. How clear are the links between data, interpretation and conclusions – i.e. how well can the route to any conclusion be seen?
15. How clear and coherent is the reporting?
16. How clear are the assumptions/theoretical perspectives/values that have shaped the form and output of the evaluation?
17. What evidence is there of attention to ethical issues?
18. How adequately has the research process been documented?

It was suggested that this framework is not prescriptive, but designed to aid the informed judgement of quality. The appraisal questions are designed to focus on four methods: in-depth interviews, focus groups, observation, and documentary analysis (Spencer, Ritchie, Lewis, & Dillon, 2003).

An alternative view on how to judge rigour and quality is offered by (Meyrick, 2006). This author generated two core principles of quality: transparency and systematicity. Figure 3 represents a quality framework centered on these principles. The framework provides a choice of techniques to establish rigour at each stage within the qualitative research process. The framework also attempts to communicate enough knowledge to enable researchers to make value judgement about rigour and quality.

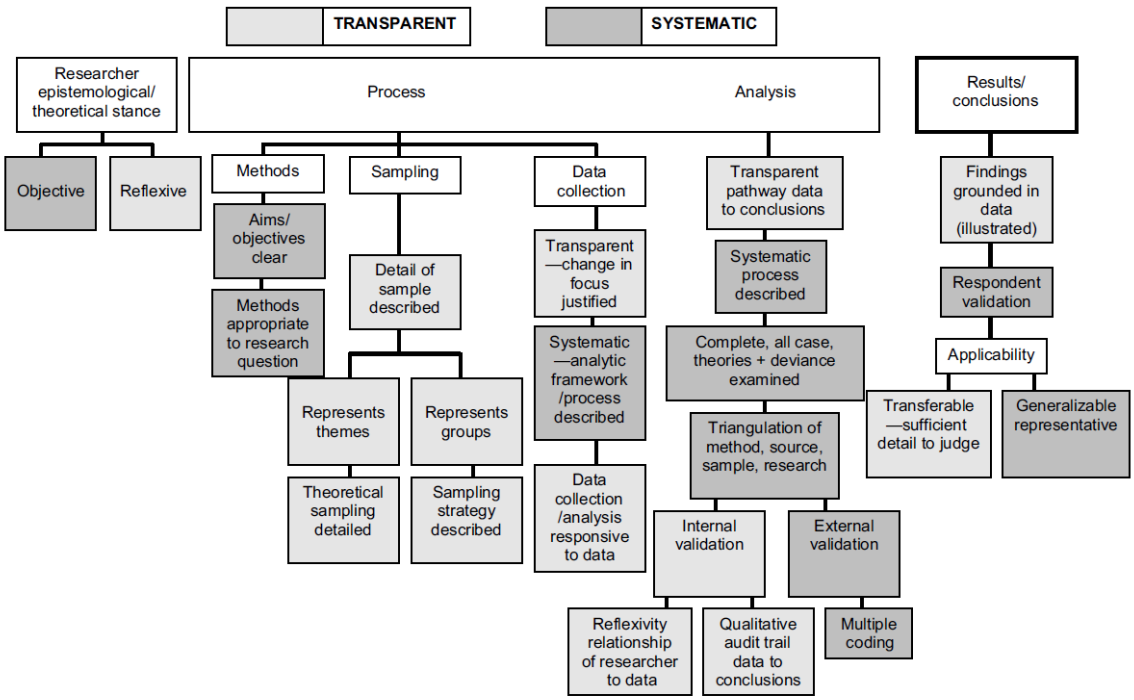

Figure 3 Quality Framework for qualitative research (Meyrick, 2006)

Discussion of various approaches to assess quality in qualitative research led to a conclusion that, in contrast to quantitative enquiries, a 'gold standard' may never be identified or achieved. Instead, the researchers should draw from approaches they consider appropriate for each individual qualitative enquiry, and apply value judgement on the basis of their knowledge and expertise.

3.7 Manual vs automated analysis of textual data

This section pinpoints research findings drawn from an individual paper presented at the conference. These findings can be considered within a broader context of research conducted by the TFF group.

A paper by (Little, Mow-Lowry, Cotton, Buick, & Blackman, 2014) described the results of a staff survey study that compared manual and automated analysis of large volume of qualitative data obtained via free-text questions to a large public sector workforce (n=102,219). The survey data generated over 180,000 distinct responses. An automated approach to analysing these data involved data analysis software Leximancer. The manual approach involved coding the data inductively to, first, reveal emerging themes and, second, supplement these themes with in-depth insights. While this research faced some methodological challenges, the major finding was that for very large datasets, an automated approach is more practical for providing useful indicators of content (themes), which can then be followed up in more depth manually, if required.

4. Discussion

This chapter contextualises the findings from the conference streams described in chapter 3 in relation to the TFF's Strategic Plan (SP) and the Code of Best Practice (CoBP). The TFF Group strives for scientific excellence in its products such as technology watch bulletins, technology forecasting and horizon scanning products, technology deep dives, and biannual emerging S&T reviews (TFF, 2014). A part of achieving excellence in scientific work is staying abreast of the latest developments in relevant scientific disciplines and research approaches. Conferences and symposia provide excellent fora for exchanging research findings in multidisciplinary settings, at national and international levels. Both emerging and state-of-the-art methods, tools and techniques are often discussed to the benefit of wider scientific community.

The ACSPRI 4th International Social Science Methodology Conference provided a forum for information exchange on a (debated) issue of quality in qualitative research. This topic is of importance to the TFF practice because much of data collected and analysed as part of the S&T Forecasting practice is of qualitative nature. The elements of qualitative research are included in all four stages of the TFF's methodology represented in Figure 4.

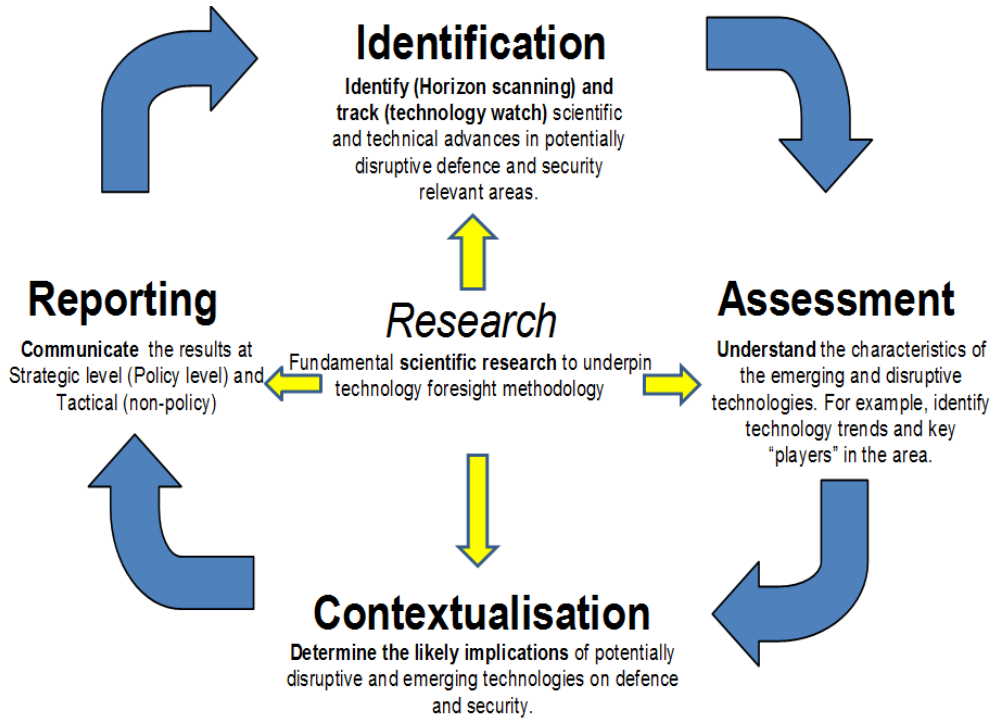

Figure 4 The TFF's overarching methodology for S&T Forecasting practice (Crone & Gaertner, 2013)

4.1 Quality and causality in qualitative research

The framework offered by (Spencer, Ritchie, Lewis, & Dillon, 2003) provides an opportunity for 'quality checks' at all stages of forecasting studies. Three out of four qualitative research methods that the appraisal questions in section 3.6 are designed for (in-depth interviews, focus groups, and documentary analysis) are used by the TFF group as part of the TFF methodology. The judgement of the quality of evidence collected by applying these methods as part of the TFF's forecasting studies can be supported by selecting appraisal questions from the list in section 3.6 on a case-by-case basis and developing quality indicators for each question according to study objectives, scope, and context.

The TFF studies can also be assessed in terms of their transparency and systematicity, as suggested by (Meyrick, 2006). The quality framework based on these principles and discussed in section 3.6 can be utilised for this purpose. It is suggested in the forecasting literature (Armstrong, Green, & Graefe, 2014) that combinations of two or more forecasting techniques that complement each other should be applied for achieving better forecasting results. By the same logic, two quality assessment frameworks discussed above can be applied, as complementary methods, to judge rigour in TFF studies. A detailed reflective research diary recommended by (Rolfe, 2006) can assist in answering the appraisal questions.

Causality in qualitative research was discussed at the conference workshop (section 3.1). At the macro level, social researchers keep looking for innovative ways to determine and highlight causality in qualitative studies, and the body of published research in this area is growing. At the micro level, the TFF's own research work is focused on developing, modifying, and applying methods that identify causal inferences in forecasting studies based primarily on eliciting expert knowledge. The Technology Development Theory (TDT) method developed by the TFF group (Mazourenko, 2015) and applied to recent studies in technology workshops is an example of this work. Technology workshops are an integral part of the TFF's Forecasting practice, particularly during the Assessment and Contextualisation phases of the methodology (Figure 4).

The QCA and Process Tracing techniques discussed at the workshop (section 3.1) may be useful tools to revise hypothesised causality in forecasting studies in medium to long term; however, the usefulness of these approaches will require further investigation.

4.2 Big Data: theory and interpretation

The Identification phase of the TFF's methodology in Figure 4 involves dealing with Big Data in order to identify and track emerging technologies and technological trends of interest to Defence. The TFF group is consistently searching for new and improved ways to detect, extract, and analyse relevant data from the torrent of Big Data.

The increasing availability of Big Data for research led to conflicting views on how to use and make sense of it (section 3.2). The view discussed and advocated at the conference was

that data-driven approaches (i.e. the view that 'data speak for themselves') underestimate: (1) the role played by researchers in the analytical process; and (2) the context necessary to identify meaningful correlations (Gonzalez-Bailon, 2013).

The scope of Big Data presentations and discussions at the conference was limited to communication dynamics and social interactions, specifically highlighting the importance of interdependence and complexity that these interactions add to social dynamics. Case studies and research findings presented at the conference provided specific examples of how Big Data can be reduced by applying filters to identify relevant streams of information, or aggregated to help identify the right temporal scale or spatial resolution. The implications related to sampling and sampling frames when collecting and analysing social data, and potential biases via sampling or filtering of Big Data were also highlighted.

It was concluded that not only social theories and models are relevant in the era of Big Data, but their role is consistently growing; and cumulative research is needed for improvements and readjustments. Furthermore, the scale of data available for analysis and analytical methods needed for successful research calls for pooling expertise from multiple disciplines. Such collaboration requires common language, which in turn demands analytical toolkits for social scientists compatible with those used in other disciplines including the mathematical language of networks, coding, and programming (Gonzalez-Bailon, 2013).

While outside the scope of the conference, the role of theory and interpretation in extracting and analysing Big Data related to forecasting studies is worth further investigation utilising the 'pointers' gained at the conference. Of particular importance are the benefits and challenges of applying multi-disciplinary analytical toolkits during the Identification phase of the TFF's methodology, and the potential sources of bias related to data collection and analysis.

4.3 Improving the quality of data: analysis of paradata

Analysing paradata *during* data collection is a technique well known to researchers who gather data via conducting structured, semi-structured and unstructured interviews (Kvale, 1996). This technique is known for its usefulness in addressing limitations of the interview design and process in real time, when an interviewer can make the necessary adjustments during fieldwork, thereby improving the quality of the dataset.

'Responsive designs' for automated data collection discussed at the conference is a relatively new, but rapidly developing field of research (section 3.3). Automated collection and analysis of paradata allow quality checks during the interview process. The aim is the same as with the 'manual' paradata collection – improving the quality of 'final' dataset – but the technical means for achieving this objective are quite different. There are computer programs and software packages developed for automated capture and processing of paradata similar to those discussed at the conference (section 3.3). These programs can be utilised for web-based surveys as well as for face-to-face or telephone interviewing.

The TFF methodology (Figure 4) includes interviewing during the Identification, Assessment and Contextualisation phases. To date, these interviews have been conducted in either face-to-face or telephone settings. While analysing paradata as part of interviewing process is common practice for the TFF group, a review of recent studies that describe how paradata were automatically analysed can be a worthy exercise, particularly for assessing and comparing 'quality indicators'.

Automated data collection can be conducted by the TFF group for studies that involve Delphi or similar techniques. In these cases, it is worth investigating how the production monitoring dashboards are organised to determine whether they are suitable for these studies given limited number of respondents (normally experts from relevant technology fields and/or ADF operators) and very 'rich' textual data.

Potentially, the TFF's forecasting practice may involve on-line surveying of larger samples. The review of similar studies that included paradata collection and analysis software would then be a worthy exercise at the study design phase.

4.4 Utilising the Total Survey Error framework

Surveys are conducted by the TFF Group as an integral part of the forecasting practice during the Identification, Assessment, and Contextualisation phases of the methodology depicted in Figure 4. Surveying limited numbers of technology experts and ADF operators is employed during data collection (primary data) and data analysis (secondary data) processes within these three phases.

The Total Survey Error (TSE) framework introduced and discussed at the conference is concerned with allocating (limited) resources is a way that maximises the quality of outputs. Since this is in accord with the TFF's own objectives, the TSE framework appears to be a useful tool that may be utilised for planning, execution, and evaluation of surveys included in TFF studies. Particularly of interest is the capacity of the TSE framework to optimise design decisions and trade-offs while reducing targeted sources of error.

The TSE framework's strengths and weaknesses listed in section 3.4 can be used for formulating research questions for investigating how this framework can be adapted for forecasting studies, and what potential benefits to the study outcomes it may bring. In addition, such investigation can be conducted within a broader context of total survey quality (TSQ). The TSQ frameworks developed and used by some organisations (section 3.4) were mentioned at the conference, but not discussed in sufficient detail. A literature review of this concept could provide suitable background for investigating the usability of the TSE framework. Of particular interest to the TFF studies are the dimensions of data accuracy, credibility, and relevance.

4.5 Addressing multi-modal on-line surveying

The increasing role of multi-modal instruments for conducting on-line surveys was highlighted at the conference. Mobile devices such as smartphones and tablets are expected to become more and more diversified in the future, with increasing rates of using these devices in comparison to stationary desk-top alternatives. The impacts on the quality of survey data collected via multi-modal devices have now been recorded and investigated (section 3.5).

A recent study conducted by the TFF Group involved using these devices for collecting rich qualitative data. The importance of addressing the implications of using mobile devices for this purpose should not be underestimated. It is imperative that these considerations are taken into account when designing TFF studies.

The published body of research in this domain is continuously growing (section 3.5). This provides an opportunity to learn from research findings and make appropriate adjustments to TFF study methodologies including a variety of formats available to study participants during data collection and data analysis stages.

4.6 Manual vs automated data analyses

The TFF studies include analyses of various types and volumes of data. At the Identification phase (Figure 4), horizon scanning and technology watch activities normally involve large volumes of data that are analysed using primarily automated means. The Assessment and Contextualisation phases include collection and analyses of small volumes, information rich data that are then analysed both automatically, utilising text analysis tools, and manually. There are trade-offs related to both ways of data analyses, and understanding the 'dynamics' of these trade-offs for the overall quality of TFF studies is an area of on-going work by the study teams.

The results of studies presented at the conference added empirical evidence from other areas of research that automated and manual approaches to data analyses can be combined. Of interest are the methodological challenges of individual studies that can be assessed and utilised as 'lessons learned' when designing TFF studies.

5. Conclusion and recommendations

The 4th International ACSPRI Social Science Methodology Conference provided a forum to a vibrant international community for exchanging study results drawn from diverse disciplines and themes in social research. A number of 'topics of interest' to the TFF Group were discussed at the conference and described in this report. Recent developments and research outcomes related to quality in qualitative research, causality, opportunities presented by Big Data, the role of paradata in increasing data quality, TSE framework,

multi-modal surveying, and the benefits of combining automated and manual data analysis provide rich material to draw from, and build on, for continuous improvements to the TFF forecasting practice.

It is recommended that:

- Quality frameworks offered by (Spencer, Ritchie, Lewis, & Dillon, 2003) and (Meyrick, 2006) are tailored to the objectives of the TFF's CoBP and assessed for their usefulness in forecasting studies
- The role of theory and interpretation in extracting 'data of interest' from the Big Data is investigated within the context of forecasting studies. Multi-disciplinary approach to this investigation can be supported by DSTO's Communities of Practice.
- The role of paradata in increasing data quality is assessed within forecasting studies. Study methodology can contain provisions for achieving this objective including the review of the relevant literature and appropriate modifications to survey design, conduct, and analysis.
- The TSE and (relevant) TSQ frameworks are assessed for their usefulness to forecasting studies. This investigation can include extensive review of the literature for making generalisations about earlier (successful) attempts to reduce survey error in similar studies followed by adapting these frameworks for 'trials' in TFF case study settings.
- The findings and recommendations of published studies related to multi-modal on-line data collection are used for designing TFF surveys
- Manual and automated data analysis techniques as complementary research mechanisms are proposed to the DSTO Communities of Practice as one of 'enduring themes' to facilitate on-going exchange between multi-disciplinary teams and learn from each other.

6. References

Ackland, R. (2014). The Social Scientist's Role in the Era of Big Data. *4th ACSPRI Social Science Methodology Conference.* University of Sydney.

Armstrong, J., Green, K., & Graefe, A. (2014). *Golder Rule of Forecasting; Be Conservative.* Working paper.

Bennett, A., & George, A. L. (1997). Process Tracing in Case Study Research. *MacArthur Foundation Workshop on Case Study Methods.*

Biemer, P. J. (2010). Total Survey Error: Design, Imlementation and Evaluation. *Public Opin Q*, 817-848.

Burt, R. S. (2011). *Structural Holes in Virtual Worlds.* Chicago: University of Chicago Booth School of Business.

Cobb, C. (2014). Dual-mode design considerations for web surveys. *ACSPRI 4th International Social Science Methodology Conference.* Sydney: ACSPRI.

Crone, D., & Gaertner, P. (2013). Toward a Systematic Process for Science and Technology Foresight. *ASOR/DORS.*

Funnell, S., & Rogers, P. (2011). Purposeful Program Theory: Effective Use of Theories of Change and Logic Models. Jossey-Bass.

Gonzalez-Bailon, S. (2013). *Social Science in the Era of Big Data.* Oxford: Oxford Internet Institute, University of Oxford.

Goodrick, D. (2014). Can qualitative approaches support causal inferences? Emerging desing options and analytic techniques. *ACSPRI Social Science Methodology Conference .*

Groves, R. M., & Lyberg, L. (2010). Total Survey Error: Past, Present, and Future. *Public Opin Q*, 849-879.

Heeringa, S. G., & Groves, R. M. (2006). *Responsive Design for Household Surveys.* University of Michigan.

Hope, K., & Waterman, H. (2003). Praiseworthy pragmatism? Validity and action research. *Journal of Advanced Nursing*, 120-127.

Kellard, K. (2014). Quality in qualitative research: what do we know? What do we need to know? *ACSPRI 4th International Social Science Methodology Conference.* Sydney: ACSPRI.

Kirgis, N. G. (2014). Using Paradata for Interviewer Data Quality Monitoring . *ACSPRI 4th International Social Scienec Methodology Conference.* Sydney: ACSPRI.

Kvale, S. (1996). *Interviews: an introduction to qualitative research interviewing.* Sage Publications.

Lattery, K., Park Bartolone, G., & Saunders, T. (2013). *Optimising Surveys for Smartphones: Maximising Response Rates While Minimising Bias.* Maritz Research.

Legewie, N. (2013). An Introduction to Applied Data Analysis with Qualitative Comparative Analysis. *Forum: Qualitative Social Research.*

Lin, J., Kelley, J., Mneimneh, Z., & Pennell, B. E. (2014). Using Paradata to Monitor Interviewer's behaviour. *ACSPRI 4th International Socia Science Methodology Conference.* Sydney: ACSPRI.

Little, G., Mow-Lowry, A., Cotton, T., Buick, F., & Blackman, D. (2014). Analysis of free text response data in large surveys: a comparison between manual and automated

analyses. *ACSPRI 4th International Social Science Methodology Conference.* Sydney: ACSPRI.

Mazourenko, E. (2015). *A Technology Development Theory Method for Science and Technoilogy Forecasting*. DSTO-TR-XXXX (forthcoming).

Meyrick, J. (2006). What is Good Qualitative Research? *Journal of Health Psychology*, 799-808.

Pennay, D. W. (2014). Introducing the Total Survey Error Framework. *ACSPRI 4th International Social Science Methodology Conference.* Sydney: ACSPRI.

Rolfe, G. (2006). Validity, trustworthiness and rigour: quality and the idea of qualitative research. *Journal of Advanced Nursing*, 304-310.

Spencer, L., Ritchie, J., Lewis, J., & Dillon, L. (2003). *Quality in Qualitative Evaluation: a framework for assessing research evidence.* Government Chief Social Researcher's Office, UK.

TFF Group. (2014). *Technology Forecasting and Futures Group 2014 Strategic Plan.* Working Draft.

Wagner, J., West, B. T., Kirgis, N., Lepkowski, J. M., Axinn, W. G., & Ndiaye, S. K. (2012). Use of Paradata in a Responsive Design Framework to Manage a Field Data Collection. *Journal of Official Statistics*, 477-499.

Whiteley, S. (2014). Mitigation Errors of Representation: a practical case study of the University Experience Survey. *ACSPRI 4th International Social Science Methodology Conference.* Sydney: ACSPRI.

DEFENCE SCIENCE AND TECHNOLOGY ORGANISATION DOCUMENT CONTROL DATA	1. DLM/CAVEAT (OF DOCUMENT)

2. TITLE	3. SECURITY CLASSIFICATION (FOR UNCLASSIFIED REPORTS THAT ARE LIMITED RELEASE USE (L) NEXT TO DOCUMENT CLASSIFICATION)
ACSPRI 2014 4th International Social Science Methodology Conference Report	Document (U) Title (U) Abstract (U)
4. AUTHOR(S)	5. CORPORATE AUTHOR
Elena Mazourenko	DSTO Defence Science and Technology Organisation PO Box 1500 Edinburgh South Australia 5111 Australia

6a. DSTO NUMBER	6b. AR NUMBER	6c. TYPE OF REPORT	7. DOCUMENT DATE
DSTO-GD-0878	AR-016-661	General Document	April 2015

8. FILE NUMBER	9. TASK NUMBER	10. TASK SPONSOR	11. NO. OF PAGES	12. NO. OF REFERENCES
			22	27

13. DSTO Publications Repository	14. RELEASE AUTHORITY
http://dspace.dsto.defence.gov.au/dspace/	Chief, Joint and Operations Analysis Division

15. SECONDARY RELEASE STATEMENT OF THIS DOCUMENT

Approved for public release

OVERSEAS ENQUIRIES OUTSIDE STATED LIMITATIONS SHOULD BE REFERRED THROUGH DOCUMENT EXCHANGE, PO BOX 1500, EDINBURGH, SA 5111

16. DELIBERATE ANNOUNCEMENT

No Limitations

17. CITATION IN OTHER DOCUMENTS Yes

18. DSTO RESEARCH LIBRARY THESAURUS

qualitative research tools; causal inferences; Total Survey Error; paradata; and multi-modal on-line surveying

19. ABSTRACT

This report summarises themes of interest to the Technology Forecasting and Futures (TFF) Group of JOAD presented at the ACSPRI 4th International Social Science Methodology conference. These themes include causal inferences in quialitative research, the role of theory and interpretation in analysing Big Data, the role of paradata in increasing data quality, the Total Survey Error framework, multi-modal on-line surveying, quality frameworks for assessing qualitative research, and the dynamics of manual vs automated data analysis. Research findings related to these themes are contextualised for the TFF S&T forecasting methodology leading to series of recommendations for improving the S&T forecasting practice.